

LAGO DISTRIBUTED NETWORK OF DATA REPOSITORIES

H. Asorey¹, A. Martínez-Méndez², L.A. Núñez^{3,4} and A. Valbuena-Delgado³ for the LAGO Collaboration⁵

We describe a set of tools, services and strategies of the Latin American Giant Observatory (LAGO) data repository network, to implement Data Accessibility, Reproducibility and Trustworthiness.

The Latin American Giant Observatory, (LAGO), is an extended continental astroparticle observatory oriented to basic research on: the Extreme Universe, Space Weather, and Atmospheric Radiation, with singles and small arrays particle detectors, covering a huge range of geomagnetic rigidity cutoffs and atmospheric absorption/reaction levels (Asorey et al. 2016).

Unlike other instruments where data flows from only one place to a network of data repositories, in LAGO each site preserves, catalogs and generates data locally which, refers not only to raw data but also to data produced during the analysis and/or simulation of cosmic rays phenomena. DSpace provides basic functionality for storing and retrieving of digital content with a straightforward adaptability for non-native types of contents and metadata schemes, supporting two interoperability protocols: OAI-PMH (Open Archive Initiatives Protocol for Metadata Harvesting) and SWORD (Simple Web-service Offering Repository Deposit) (Smith et al. 2003; Lewis et al. 2012). We have overcome one of the most important DSpace limitations: its inability to upload/download multiple records, developing a script to ingest data profiting from the some DSpace capabilities (Asorey et al. 2015). This data repository network will also be useful for the solar physics and space climatology communities.

The Data Accessibility, Reproducibility and Trustworthiness (DART) initiative was launched by CHAIN-REDS (Coordination and Harmonisation of Advanced e-infrastructure for Research and Educa-



Fig. 1. Continental LAGO data repository network.

tion Data Sharing), an European Commission co-funded project focused on promoting and supporting technological and scientific collaboration across different communities in various continents. GRNET PID (European Persistent Identifiers Consortium) service enables the allocation, management and resolution of PIDs and has been employed to ensure the data persistence and reproducibility of the experiments. Part identifiers can compute an unlimited number of handles on the fly, without requiring registering each separately. (Barbera et al. 2014).

This project has been partially funded by VIE Universidad Industrial de Santander.

REFERENCES

- Asorey, H., Cazar-Ramírez, D., Mayo-García, R., Núñez, L., Rodríguez-Pascual, M., Torres-Niño, L., & the LAGO Collaboration. 2015, in *The 34th International Cosmic Ray Conference*, Vol. PoS(ICRC2015), 672
- Asorey, H., Núñez, L., Suárez-Durán, M., Torres-Niño, L., Rodríguez-Pascual, M., Rubio-Montero, A., & Mayo-García, R. 2016, in *Cluster, Cloud and Grid Computing (CCGrid)*, 2016 16th IEEE/ACM International Symposium on, IEEE, 707–711
- Barbera, R., Becker, B., Carrubba, C., Inserra, G., Jalife-Villalón, S., Kanellopoulos, C., Koumantaros, K., Mayo-García, R., Núñez, L., Prnjate, O., Ricceri, R., Rodríguez-Pascual, M., Rubio-Montero, A., Ruggieri, F., & Project, C.-R. 2014, in *ANAIS DAS SESSÕES TEMÁTICAS E PÔSTERS*, 166
- Lewis, S., de Castro, P., & Jones, R. 2012, *D-Lib Magazine*, 18
- Smith, M., Barton, M., Bass, M., Branschovsky, M., McClellan, G., Stuve, D., Tansley, R., & Walker, J. H. 2003, *D-lib magazine*, 9
- ¹Laboratorio Detección de Partículas y Radiación, Centro Atómico Bariloche & Instituto Balseiro, Bariloche, Argentina
- ²Escuela de Ingeniería de Sistemas, Universidad Industrial de Santander, Bucaramanga, Colombia.
- ³Escuela de Física, Universidad Industrial de Santander, Bucaramanga, Colombia.
- ⁴Departamento de Física, Universidad de Los Andes, Mérida, Venezuela.
- ⁵The Latin American Giant Observatory (LAGO), <http://lagoproject.org>, see the full list of members and institutions at <http://lagoproject.org/collab.html>